Development of Annual Low-flow Regional Regression Equations for Idaho

Jon Hortness
Hydraulic Engineer
USGS, Boise

USGS
science for a changing world

Presentation Points

- Background
- Need for low-flow equations
- Equation development procedures
- Using regression equations today

Regression Equation Background

- Used to estimate streamflow statistics at ungaged locations
- Most common are the “Peak-flow” equations
  - Estimate magnitude of flood peaks at specific recurrence intervals

Status of Idaho Equations

- Monthly exceedance/annual mean equations (developed)
  - 20-, 50-, and 80-percent exceedances for each month and annual mean flow
  - Hortness and Berenbrock, 2001; USGS WRIR 01-4093
- Peak-flow equations (updated)
  - 2-, 5-,10-, 25-, 50-, 100-, 200-, and 500-year recurrence intervals
  - Berenbrock, 2002; USGS WRIR 02-4170

Status… – cont.

- Bankfull flow equations (developed)
  - 1.5- and 2.33-year recurrence intervals
  - Hortness and Berenbrock, 2003; USGS WRIR 03-4261

Annual Low-flow Equations

- Used for water quantity and/or water quality standards
  - River management planning
  - Permitting processes for withdrawals, transfers, or effluent discharges
  - Minimum thresholds for aquatic biota
  - Land-use planning and regulation

Why Now???

- Idaho TMDL's
  - Many TMDL's require streamflow data to calculate loads
  - Streamflow data is needed at locations where current data is not available
- Intermittent vs. Perennial Streams
  - Idaho Administrative Code, Water Quality Standards:

    Intermittent streamflow --- $7Q_2 < 0.1$ ft^3/s

Current Project

- Two-year study funded jointly by IDEQ and USGS
  - Develop the following annual low-flow regression equations:
    - $7Q_1$
    - $7Q_2$
    - $7Q_{10}$
    - $30Q_5$
  - Better determine locations of intermittent and perennial stream reaches

Development Process

Regionalization

- Divide state into regions
  - Try to define regions that are "hydrologically homogeneous"
  - Cluster analyses were performed on specific basin characteristics of gaged sites
    - Mean basin elevation
    - Basin slope
    - Percent forest cover
    - Mean annual precipitation

Regionalization - cont.

- Criteria for final region boundaries:
  - Cluster analyses
  - Geographic features (mountain ranges, breaks between mountains and plains, etc.)
  - Engineering judgment (general knowledge of areas)

  *** This is a subjective process!!!
- Monthly exceedance/annual mean equations
  - 8 regions and 1 undefined area
- Peak-flow and bankfull equations
  - 9 regions and 1 undefined area

Regions

- Monthly exceedance and annual mean flow equations

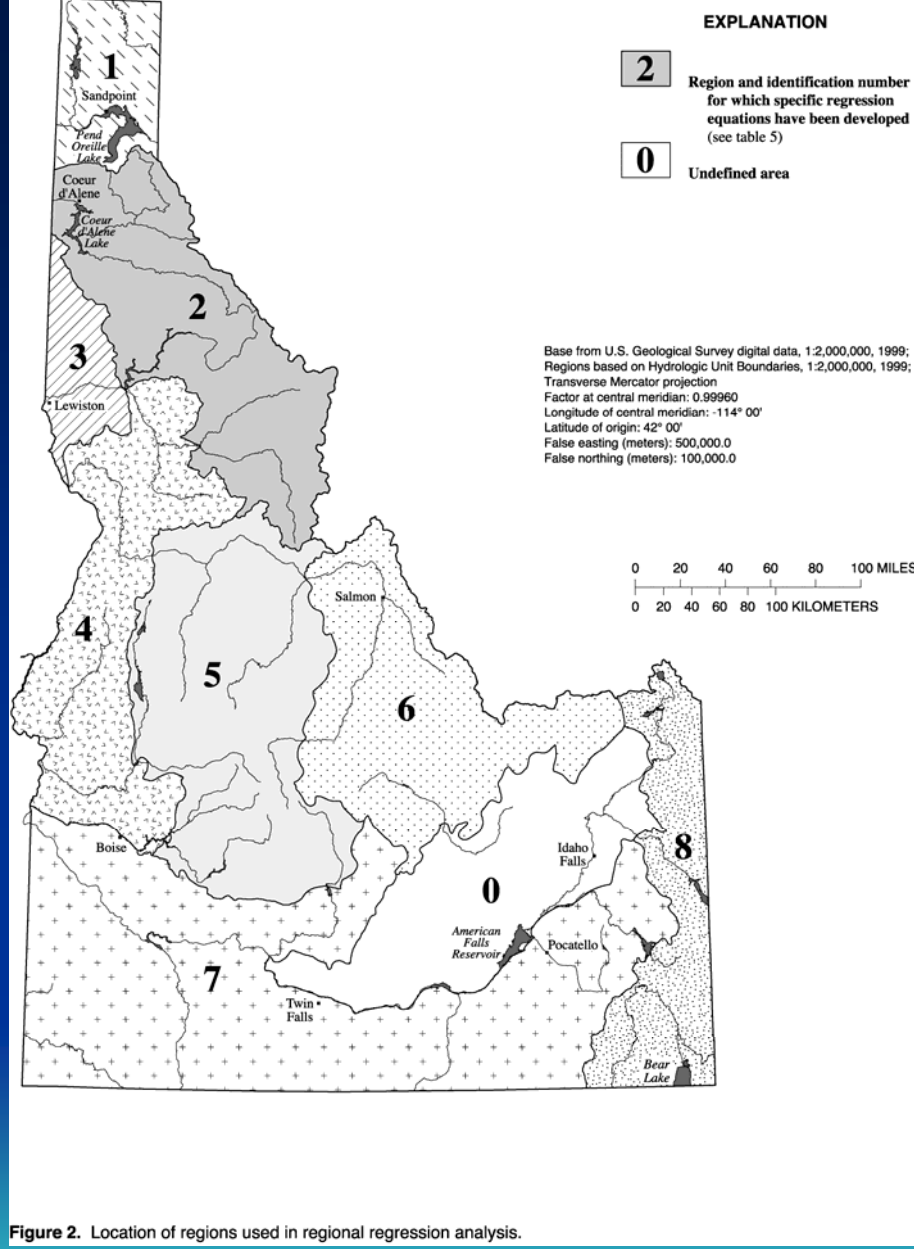


Figure 2. Location of regions used in regional regression analysis.

USGS science for a changing world

Regions

- Peak-flow and bankfull equations

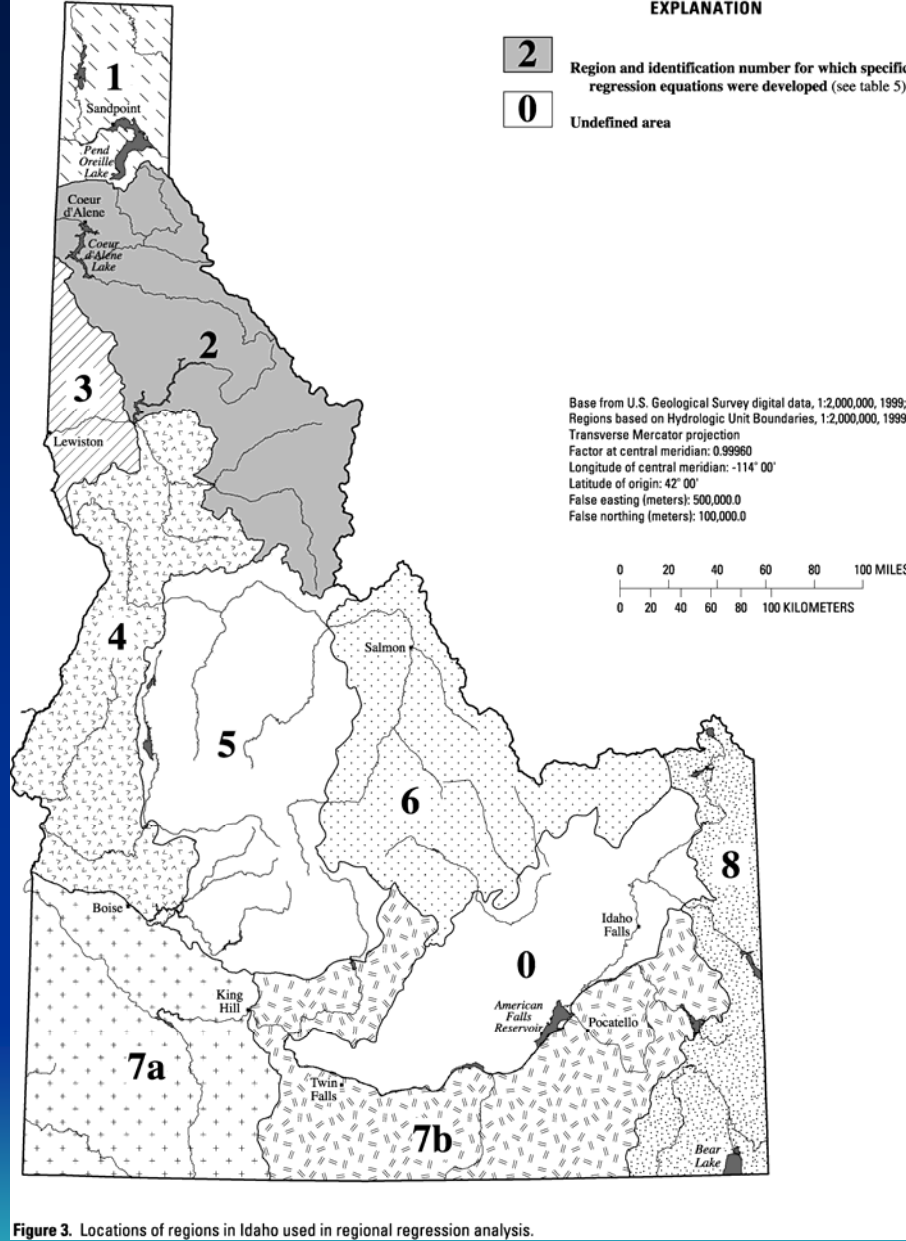


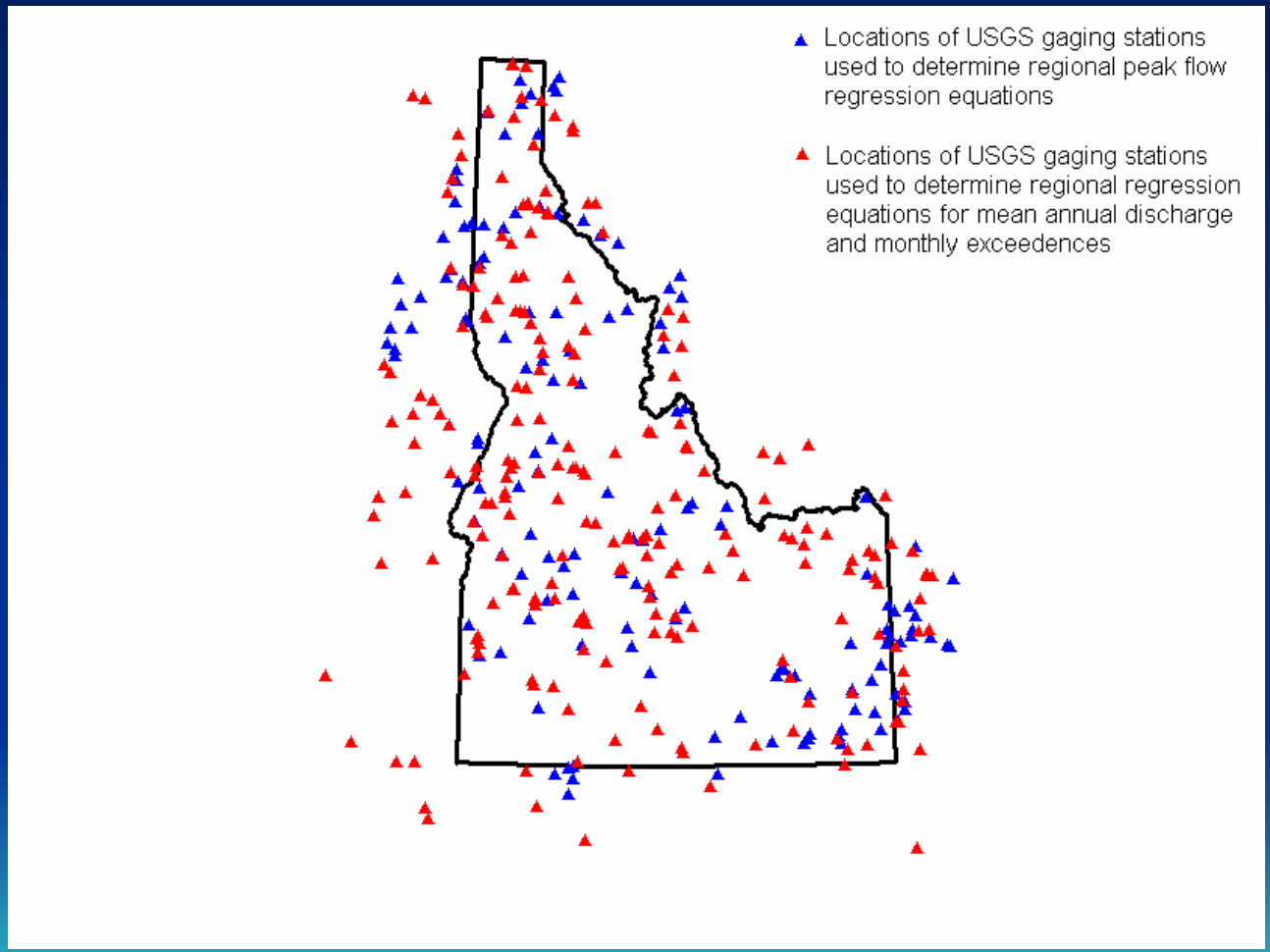
Figure 3. Locations of regions in Idaho used in regional regression analysis.

USGS science for a changing world

Gaging Station Data

- Streamflow statistics from active and discontinued gaging stations are used to develop the regression equations
  - Data from gages with 10 or more years of record on unregulated streams
  - 333 gages used in the peak-flow and bankfull flow analyses
  - 200 gages used in the monthly exceedance and mean annual flow analyses

Gaging Station Locations



USGS
science for a changing world

Basin and Climatic Characteristics

- Obtained from digital data sources using GIS applications
  - Differ from previous investigations --- all basin and climatic characteristic data are computer generated and can be easily reproduced
    - No differing of estimates because two people calculate the same characteristic two different ways
  - Allows for use of characteristics that were previously very difficult or impossible to determine manually

Basin and Climatic Characteristics – cont.

- Characteristics used in the recently updated/developed equations
  - Drainage area
  - Mean basin elevation
  - Basin relief
  - Slopes > 30%
  - North-facing slopes > 30%
  - Mean annual precipitation
  - Forested area
  - Basin slope
  - Main channel slope
- List is almost endless
  - If the data are available, it's probably possible to use it

Regression Analyses

- Multiple linear regression techniques
  - Regress a specific streamflow statistic for all gages within a region vs. basin and climatic characteristics for the same gages
  - Use an OLS (Ordinary Least Squares) stepwise analyses to determine the best possible combination(s) of characteristics
    - The following regression statistics are useful in determining the “best” final equation
      - Student’s t; Cook’s D; Mallow’s C_p; Adjusted R^2; Press

Regression Analyses - cont.

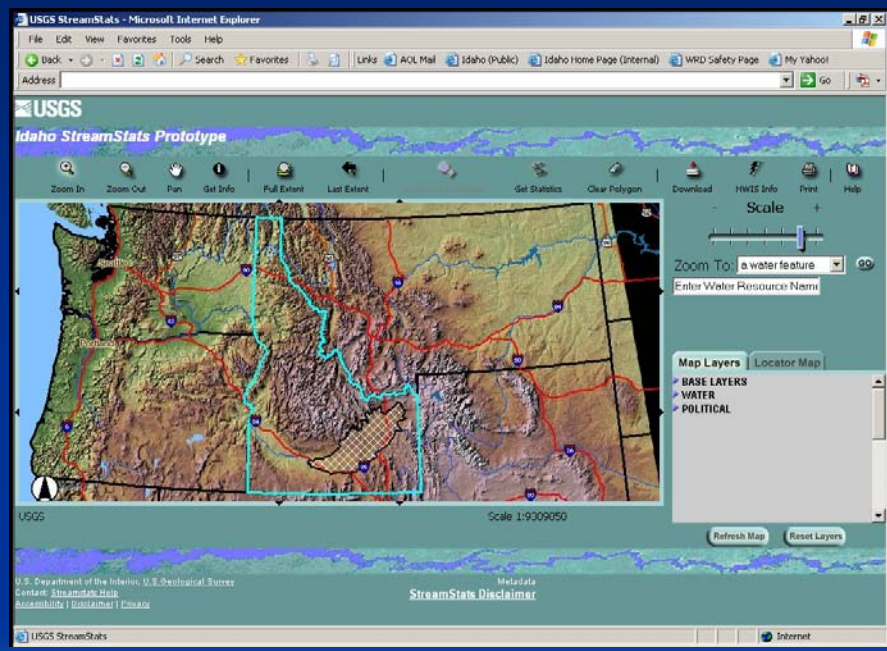
- Use a GLS (Generalized Least Squares) technique, developed by USGS, to finalize the equations
  - Often reduces the prediction error because it weights the data from each gage based on record length and standard deviation of the data
- The final equation is not always the one that produces the smallest prediction errors
  - Number of variables (don't "over-fit" the data)
  - Ease of use
- Example equation: $Q_{100} = 0.0607\ DA^{0.775}\ P^{1.96}$

Using Regression Equations in 2005

- Use of digital datasets
- Computer generated basin and climatic characteristics (GIS-based)
- More robust regression techniques

Using Regression Equations in 2005 – cont.

- StreamStats
  - Web-based, interactive map for using regression equations to estimate streamflow statistics
  - Prototype version for Idaho (beta testing)
  - Will be released to the public this year
  - Hope to implement it nationwide as new equations are developed



- Questions:
  - Al Rea, Idaho District, USGS
  - 387-1323; ahrea@usgs.gov

USGS science for a changing world

QUESTIONS???

Jon Hortness
387-1319
jehortne@usgs.gov